1. (15 points) Let a_1 and a_2 be columns of matrix a. Let (x_i b_{ji}) be paired data, where i runs from 1 to n.

$$E(\mathbf{a}) = \frac{1}{2n}\sum_{i=1}^{n}\sum_{j=1}^{2}(\mathbf{x}_i^T\mathbf{a}_j - b_{ji})^2$$

Derive $\frac{d}{d\mathbf{a}}E(\mathbf{a})$

2. (15 points)Let $E(\mathbf{a}) = \frac{1}{2n}\sum_{i=1}^{n}\sum_{j=1}^{2}(\tanh(\mathbf{x}_i^T\mathbf{a}_j) - b_{ji})^2$

Derive $\frac{d}{d\mathbf{a}}E(\mathbf{a})$

3. (20 points) Draw a flow chart to illustrate minimizing E(a) of problem 2 by the gradient descent method.

4. (25 points) Write function gradient_descent to estimate the matrix for linear transformation. Checked by _____________________time_____________

```
x=rand(400,2);
z(:,1) = 2*x(:,1)+x(:,2)-1;
z(:,2)=x(:,1)-x(:,2)+1;
a=gradient_descent(x,z)
```

5. (25 points) Write function gradient_descent to estimate the matrix for nonlinear transformation. Checked by ____________________time_________________

```
x=rand(400,2);
z(:,1) = tanh(2*x(:,1)+x(:,2)-1);
z(:,2) = tanh(x(:,1)-x(:,2)+1);
a=gradient_descent(x,z)
```